Gene expression physiology and pathophysiology of the immune system

Louis M. Staudt

Genomic-scale gene expression profiling can reveal cellular physiology with unprecedented richness. This technology is being used to define the gene expression targets of individual regulatory proteins and signaling pathways. Comprehensive databases of gene expression measurements can be used to understand the pathological mechanisms underlying disease processes.

Louis M. Staudt
Metabolism Branch, NCI Metabolism Branch, Division of Clinical Sciences, National Cancer Institute, Bldg 10, Rm 4N114, Bethesda, MD 20892, USA.
e-mail: lstaudt@box-l.nih.gov

Journals and websites are serving up genomic data at an exponentially increasing rate and biology will have to adapt. An appreciation for biological systems in aggregate will have to develop. Genomic data will need to be organized, synthesized and modeled to form a physiological description of the system under study. This review will focus on the emerging field of 'gene expression physiology', which attempts to understand the regulation of mRNA abundance on a genomic scale. One goal of gene expression physiology is to understand which genes are downstream transcriptional targets of particular cellular signaling pathways (Fig. 1). Using recently developed high-throughput technologies, a comprehensive gene expression 'profile' of a cell can be readily obtained. Gene expression profiling can be used, for example, to define the changes in cellular physiology that occur as an immune cell is stimulated by a cytokine or by antigen. Likewise, gene expression profiling can read out the transcriptional consequences of perturbing an individual signaling pathway. Experimentally, this can be achieved by overexpressing a wild-type or dominant negative version of a regulatory protein in the pathway, by disrupting a gene regulating the pathway, or by treating cells with pharmacological inhibitors of the pathway. The use of gene expression profiles collected from well-defined experimental systems is amplified as they are assembled into large gene expression databases. When the gene expression profile of a disease process is subsequently obtained, the pre-assembled gene expression database can be used to gain insight into the signaling pathways that are pathologically engaged in the disease.

Microarray analysis of gene expression physiology

The study of gene expression physiology has been jump-started by new technologies that quantitate, in parallel, the mRNA levels of tens of thousands of genes[1–3]. The two dominant technologies, spotted cDNA microarrays and oligonucleotide arrays, each begin with an ordered array of nucleic acids representing thousands of genes on a solid support[4]. mRNA from the cells of interest is used to create a fluorescent, first-strand cDNA probe that is then hybridized to the microarray. In many cases, especially when using spotted cDNA microarrays, two mRNA samples are directly compared on the same microarray by incorporating different fluorochromes into the cDNA probes derived from the samples. The extent of hybridization of the probes to each gene on the microarray is then quantitated and the ratio of the pixel intensities for each fluorochrome is an excellent measure of relative gene expression in each mRNA sample. When microarray experiments are of high technical quality, the derived fluorescence ratios are in close quantitative agreement with relative gene expression measurements derived from Northern blots or quantitative RT-PCR.

The relative ease and robustness of DNA microarray methodology has led to an explosion of gene expression information in the past few years. To harness this data stream, a diversity of mathematical methods have been developed that all aim at finding the coherent patterns inherent in gene expression data sets[5]. One of the most popular methods, hierarchical clustering, uses a standard Pearson correlation coefficient to quantitate the similarity in expression of two different genes across a set of samples[6]. Using this similarity measure, a list of genes can be ordered hierarchically, leading to an appreciation of sets of genes that are coregulated. The same methods can be used to determine the similarity in gene expression between different samples. Often the results of these mathematical manipulations are displayed in a tabular format (Fig. 2) in which highly expressed genes are depicted in shades of red and underexpressed genes are depicted in shades of green[6]. Each row in such a table represents the measurements for a given gene across several mRNA samples, and each column represents the results from a single sample. This visualization tool often reveals the gene expression physiology inherent in a complex DNA microarray data set.

Systematic gene expression analysis of the immune system

The announced completion of the human genome sequence is only the beginning of a complete description of human genes. A daunting task that lies

http://immunology.trends.com 1471–4906/01/$ – see front matter PII: S1471-4906(00)01792-0

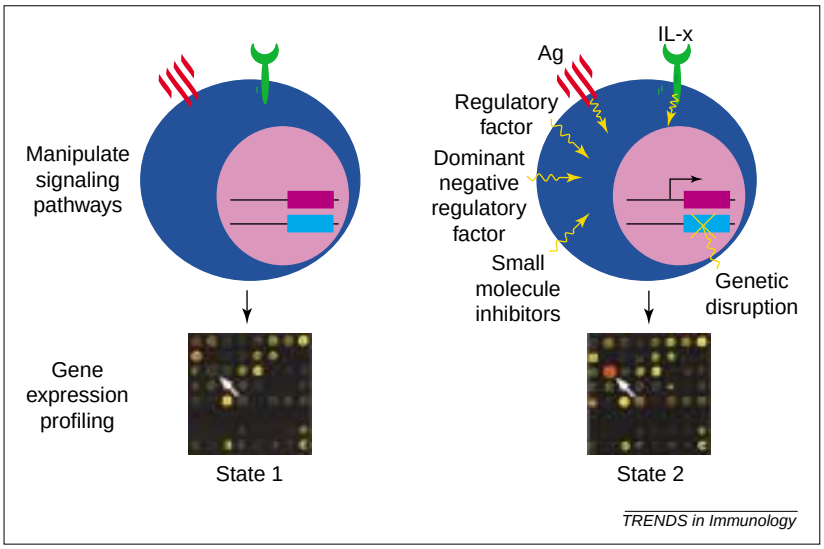

Fig. 1. Use of gene expression profiling and signaling pathway manipulation to elucidate the gene expression physiology of a cell.

ahead is to completely annotate this genomic sequence to describe all human genes that are expressed as mRNA. When this is accomplished, it will be possible to create a single microarray that represents all human genes. In the interim, it has proven very useful to create specialized cDNA microarrays that focus on genes of known or suspected importance to a particular biological system. One such microarray, the lymphochip, was created to analyze normal and disease states of the immune system[7,8]. The starting point of the lymphochip was the high-throughput sequencing of cDNA clones from libraries derived from human immune cells. The expressed sequence tags (ESTs) derived from this effort were mined for those representing novel genes that were enriched in immune cell cDNA libraries. The lymphochip incorporated 15 000 such clones from libraries prepared from germinal center B cells as well as from various lymphoma and leukemia subtypes. In addition, roughly 3500 'named' genes of known structure and function were included on the lymphochip based on their published roles in immune cell differentiation, responses and disorders.

Figure 2 presents over one million measurements of gene expression from 202 lymphochip analyses of gene expression in normal and pathological immune cells. The genes in this table were organized using the hierarchical clustering algorithm described above, revealing groups of genes that were over- or under-expressed in groups of samples. One of the powerful messages that has emerged from genomic-scale analysis of the simple eukaryote *Saccharomyces cerivisiae* is that genes that function in similar biochemical or functional pathways are coordinately regulated across a variety of cellular states. For example, the diauxic shift of yeast from anaerobic to aerobic metabolism is accompanied by the coordinate induction and repression of genes in the glycolytic pathway[9]. Likewise, genes encoding subunits of the same multi-protein complex (for example, the proteosome and ribosome) are often strikingly coregulated[6].

Similar coregulation of functionally related genes has been observed in mammalian cells in a number of large gene expression profiling experiments[7,10,11]. Genes that are coordinately expressed in a particular cell type or during a biological response have been termed gene expression 'signatures'[7]. Some of the gene expression signatures listed at the left of Fig. 2 highlight genes expressed at certain stages of lymphoid differentiation (for example, the pan-B-cell and germinal center B-cell signatures). The proliferation signature is defined by hundreds of genes that are coordinately regulated during the cell cycle, demonstrating that much of the gene expression physiology of the cell is tied to this critical biological response. Other signatures represent induced gene expression during activation of immune cells by mitogens and/or cytokines (for example, the lymphocyte/monocyte activation and T-cell activation signatures) or response of these cells to insults (for example, the DNA damage signature). Finally, some signatures represent pathological gene expression in immune disorders and malignancies [for example, signatures of rheumatoid arthritis synovial cells (RA), chronic lymphocytic leukemia (CLL) and multiple myeloma]. Representative genes from these signatures are shown at the right of Fig. 2, but each of these signatures includes scores of genes, many of unknown function. Such a gene expression map is therefore a powerful hypothesis generator that can provide clues about which genes are functionally involved in particular cellular responses.

Recent gene expression profiling experiments have yielded gene expression signatures of B-cell activation and tolerance[12], T-cell activation[13], and T-helper cell differentiation[14]. B-cell tolerance induced by chronic exposure to self antigen generated a gene expression signature that was surprisingly unrelated to normal lymphocyte activation responses[12]. Genes upregulated by antigen stimulation of normal B cells were not highly expressed in tolerant cells, including genes that are required for mitogenesis (*c-myc* and *LSIRF/IRF-4*) and for protection from activation-induced apoptosis (A1). Further, tolerant B cells expressed LKLF, a negative regulator of lymphocyte proliferation that is normally downregulated by BCR signaling in normal B cells. CD72 was found to be upregulated in tolerant cells, which might inhibit BCR signaling in tolerant cells by recruitment of SHP-1. Thus, the gene expression signature of tolerant B cells explained, in large measure, the anergic phenotype of these cells.

Two subtypes of helper T cells, Th1 and Th2, secrete distinct spectra of cytokines and participate differentially in various immune-mediated diseases. The selective homing of Th1 versus Th2 cells to sites of inflammation may be understood based on the gene expression signatures of these two cell types[14]. First, Th1 cells treated with IL-12 upregulated fucosyltransferase VII (FUT7), which is responsible for fucosylating selectin ligands on the surface of T cells.

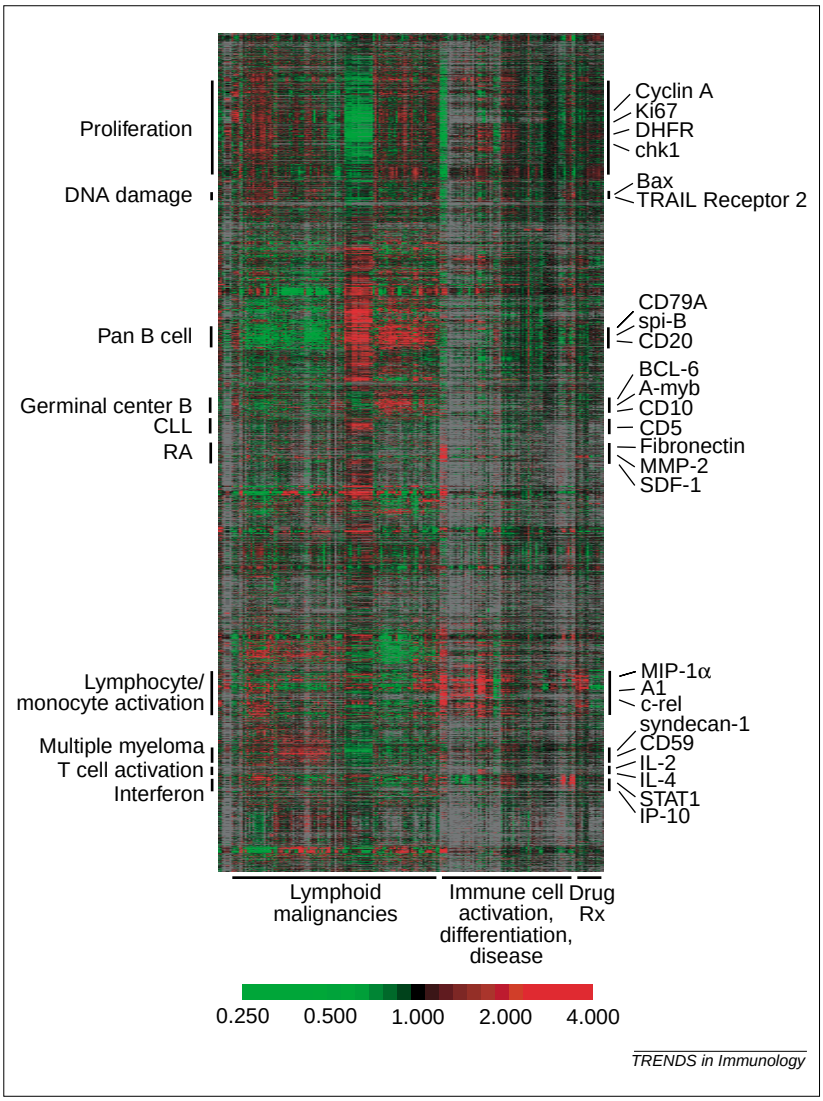

Fig. 2. A gene expression map of normal and pathological immune responses.The table presents lymphochip microarray data from 202 experiments analyzing gene expression in lymphoma, leukemia, and multiple myeloma cell lines and samples, in autoimmune diseases, in immune cells stimulated with various mitogens and cytokines, and in cell lines treated with pharmacological inhibitors.The genes were organized by hierarchical clustering (see text for details). Gene expression data are depicted over a 16-fold range using the color scale shown at the bottom. Grey represents bad or missing data. Gene expression signatures consisting of coregulated genes are shown on the left; representative genes in each gene expression signature are shown on the right.

This fucosylation is required for the first step of lymphocyte adhesion to endothelial cells, 'rolling'. IL-12 also induced two chemokine receptors, CCR5 and CCR1, both of which promote increased interaction between lymphocytes and chemokines (RANTES and MIP-1α) bound to the surface of endothelial cells. Finally, IL-12 upregulation of integrin α6 promotes binding of Th1 cells to laminin, leading to successful diapedesis and entry into the inflammatory site.

Subtyping of diseases by gene expression profiling

One of the powerful applications of gene expression profiling to human disease relies on its ability to provide a rich and quantitative molecular phenotype of the disease process. For diseases in which the diagnosis relies on clinical observations together with non-quantitative assays such as histological morphology, gene expression profiling has the potential to subdivide current diagnoses into molecularly distinct diseases. Genomic-scale gene expression profiling of lymphomas and leukemias has provided proof-in-principle of this analytic paradigm[7]. Known subtypes of these malignancies could be readily distinguished from each other by their characteristic gene expression signatures[7,15]. Chronic lymphocytic leukemia, the most common leukemia of adults, had a gene expression signature that was shared, in part, by resting blood B cells[7], a finding that is consistent with the indolent, relatively non-proliferative character of this disease. By contrast, follicular lymphoma cells strongly resembled normal germinal center B cells in gene expression[7]. This finding is in keeping with the observation that follicular lymphoma cells show evidence of ongoing somatic mutation of immunoglobulin genes[16], a process that is characteristic of germinal center B cells.

Diffuse large B-cell lymphoma (DLBCL), the most common subtype of non-Hodgkin's lymphoma, has long been a clinical and pathological enigma. Although various classification schemes have subdivided this diagnostic category on morphological grounds, these subdivisions could not always be reproducibly applied. Clinically, it has been puzzling that combination chemotherapy is able to cure approximately 40% of these patients while failing in the remainder. A clear possibility, therefore, was that this diagnostic category actually lumps together more than one molecularly distinct disease, a suspicion borne out by gene expression profiling. One subtype of DLBCL, termed germinal center B-like DLBCL, expressed most genes in the germinal center B cell signature and thus was putatively derived from this stage of B cell differentiation (Fig. 3)[7]. The other subtype of DLBCL, termed activated B-like DLBCL, lacks expression of these germinal center B cell genes and instead expresses genes that are induced in blood B cells by mitogenic stimulation. In keeping with this subdivision of the DLBCLs was the observation that the germinal center B-like lymphomas had evidence of ongoing somatic mutation of immunoglobulin genes, whereas the activated B-like DLBCLs did not[17]. Clinically, these two DLBCL subsets were remarkably divergent: 75% of the germinal center B-like DLBCL patients were alive five years following chemotherapy as compared with less than one quarter of patients with activated B-like DLBCL (Ref. 7). This example highlights the virtue of genomic-scale gene expression profiling of normal lymphocyte subsets in order to provide a framework to interpret gene expression profiles of pathological processes.

What other immune disorders might benefit from the diagnostic precision of gene expression profiling? Most of the autoimmune disorders are heterogeneous in clinical presentation and course. The diagnosis of systemic lupus erythematosis (SLE), for example, requires some, but not all, clinical features of the

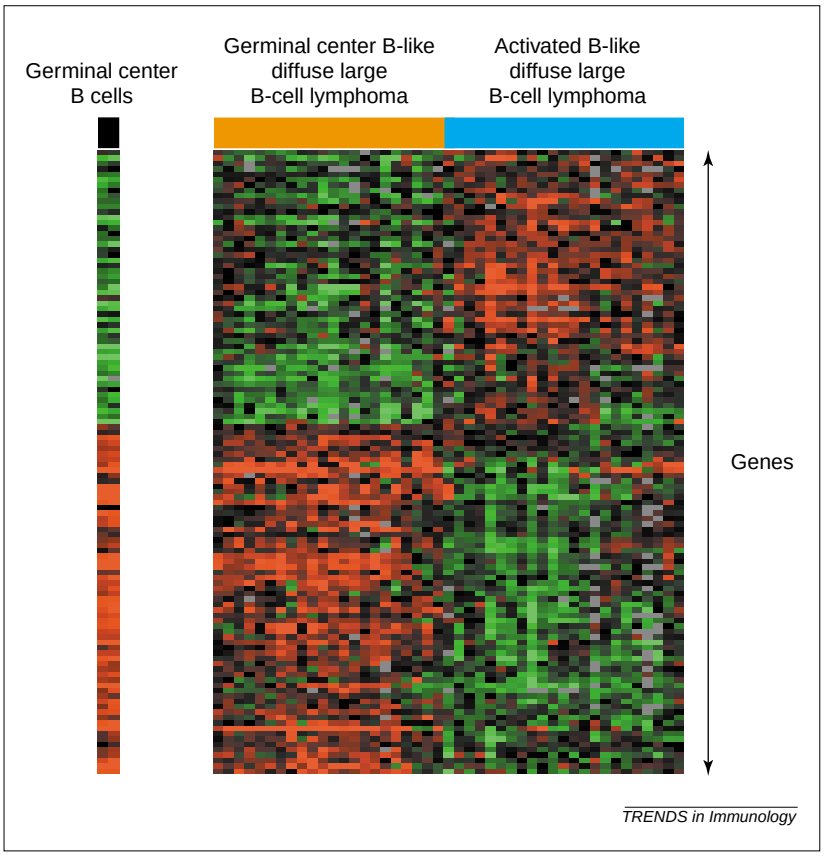

Fig. 3. Discovery of two types of diffuse large B cell lymphoma (DLBCL) by gene expression profiling. Genes were selected that highly correlate with the distinction between germinal center B-like DLBCL and activated B-like DLBCL (Ref. 7). Expression of these genes in samples of DLBCL and in tonsillar germinal center B cells is shown according to the color scale shown in Fig. 2.

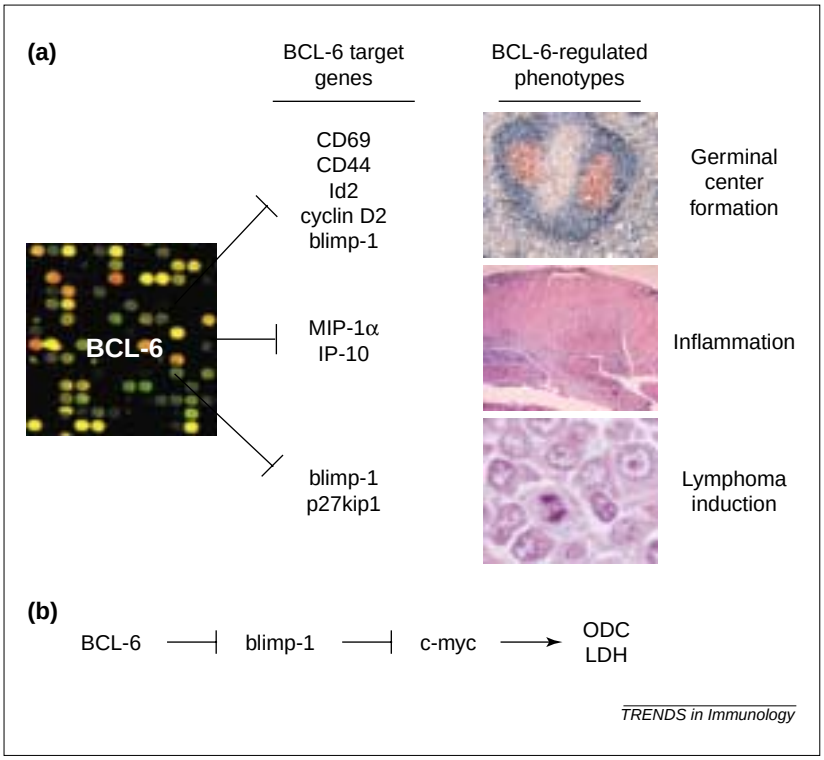

Fig. 4. DNA microarray analysis of BCL-6 target genes. (a)Target genes that are repressed by BCL-6 shown together with the BCL-6 regulated phenotypes that they help to explain. See text for details. (b) A regulatory cascade of transcription factors initiated by BCL-6.

disease to be present and shows evidence of polygenic inheritance[18,19], suggesting that SLE may be a collection of related, but molecularly distinct diseases. Heterogeneity in disease definition will certainly confound attempts to map genes responsible for SLE. Conceivably, each of the underlying genetic loci that contribute to SLE might confer a characteristic gene expression signature that could be used as a quantitative marker to map the gene in family studies of SLE. Molecular heterogeneity in other autoimmune diseases might also explain the variable responses of patients to therapy. The challenge facing those who wish to profile gene expression in immune-mediated diseases, however, is to choose the appropriate source of cells to profile. One ready source is peripheral blood, although it is not yet clear that the pathogenic cell types in these disorders will constitute a large enough proportion of blood cells to be detected by gene expression profiling. Biopsies of sites of active disease might be helpful, but it may be necessary to fractionate the biopsied cells into defined cellular subtypes prior to gene expression profiling.

Dissecting the molecular circuitry of the cell

The applications of gene expression profiling mentioned thus far take particular advantage of the genomic-scale portrait of gene expression that this technology provides. A complementary use of DNA microarrays is to survey the changes in gene expression that occur when specific regulatory factors or pathways are perturbed. Such perturbations could be experimentally effected by dominant gain- or loss-of-function mutations in regulatory factors, by ablation of regulatory genes in the germ line or in somatic cells, or by pharmacological manipulations. A simple model of gene expression physiology would suggest that activation of a particular signal transduction pathway would lead to a characteristic and limited set of gene expression changes. If this were the case, then each signal transduction pathway could be investigated one by one, leading eventually to a model that could predict the gene expression fluxes that would result from changes in a cell's environmental stimuli. Although this is clearly an oversimplified view, recent experiments provide tempered optimism that this bottom-up approach to gene expression physiology will be fruitful.

Transcription factors are attractive starting points for this approach since their site of action is proximal to the gene and thus their action might be less confounded by regulatory cross-talk than more upstream components of signal transduction pathways. Overexpression and/or dominant negative methodology has been coupled to DNA microarray technology to define the regulatory targets of the transcription factors p53 (Ref. 20), c-myc (Refs 21,22), wt1 (Ref. 23), Pax3–FKHR fusion oncoprotein[24], and BCL-6 (Ref. 25). It is no accident that all of these transcription factors are oncogenes or tumor suppressors: translocation, overexpression or loss of

these factors presumably causes cancer because they have non-redundant and/or limiting effects on gene transcription.

DNA microarray analysis of BCL-6 target genes

To illustrate this type of microarray analysis, Fig. 4 summarizes the target genes of BCL-6, a frequently translocated oncogene in diffuse large B-cell lymphoma that also regulates normal immune responses[26,27]. BCL-6 mutant mice have a variety of immune system defects including inability to form germinal centers and a spontaneous and lethal inflammatory disease of the heart and lungs that is characterized by Th2-type cytokines[28–30]. BCL-6 protein is expressed highly in most germinal center B cells[31–33] and the defect in germinal center formation in BCL-6 mutant mice is intrinsic to B cells[30]. Because BCL-6 is a potent transcriptional repressor, lymphochip microarrays were used to screen for genes that were downregulated after introduction of wild-type BCL-6 into B cell lines that lack endogenous BCL-6 expression[25]. In addition, a dominant negative form of BCL-6 consisting only of its DNA-binding domain was introduced into B cell lines that naturally express BCL-6, and lymphochip microarrays were used to detect genes that were 'de-repressed'. Only genes that were affected by both of these manipulations were deemed presumptive BCL-6 target genes.

The BCL-6 target genes provided rich insights into the known BCL-6-regulated phenotypes (Fig. 4a). One set of BCL-6 target genes consists of many B cell activation genes such as CD69, CD44, Id2 and cyclin D2. These genes are induced by BCR stimulation of resting B cells and BCL-6 was able to block this induction[25]. By contrast, none of these genes is expressed in germinal center B cells, which have high levels of BCL-6 protein. A second important BCL-6 target gene was blimp-1, a transcriptional repressor that is critical for plasmacytic differentiation[34–36]. The ability of BCL-6 to repress blimp-1 suggests that BCL-6 blocks plasmacytic differentiation. Indeed, dominant negative BCL-6 was able to cause partial plasmacytic differentiation in a Burkitt's lymphoma cell line[25]. These target genes might explain how BCL-6 controls the fate of a B cell following antigen exposure. A naive splenic B cell encountering antigen can either become activated and differentiate rapidly into plasmacytic cells in the periarteriolar lymphoid sheath or differentiate into a germinal center B cell in the follicular region[37]. By blocking expression of B cell activation genes and blimp-1, BCL-6 might skew the fate of the B cell towards the germinal center program.

The inflammatory phenotype of BCL-6 mutant mice may be explained, in part, by the fact that BCL-6 inhibits expression of the chemokines, MIP-1α and IP-10 (Ref. 25). These chemokines attract monocytes and activated T cells to sites of inflammation, and their derepression in BCL-6 mutant mice could contribute to the observed myocarditis and pulmonary vasculitis.

Finally, BCL-6 target genes provide a plausible mechanism by which BCL-6 causes lymphomas. Differentiation of germinal center B cells into plasma cells is accompanied by loss of BCL-6 expression[30], thus allowing blimp-1 to be expressed. Translocation of BCL-6 in non-Hodgkin's lymphomas prevents this physiological downregulation of BCL-6 expression, thus blocking blimp-1 expression and trapping the cell at the germinal center stage of differentiation. A target gene of blimp-1 repression is c-myc and, therefore, BCL-6 translocations would indirectly maintain progression through the cell cycle by elevating c-myc expression. These findings place BCL-6 at the top of a regulatory cascade of transcription factors (Fig. 4b). Cell-cycle progression is also promoted by the ability of BCL-6 to repress p27kip1, a cyclin-dependent kinase inhibitor[25]. Thus, BCL-6 translocations co-opt the normal regulatory functions of BCL-6, thereby promoting proliferation, preventing terminal plasmacytic differentiation and possibly allowing secondary oncogenic hits to further transform the cells.

Will this straightforward approach uncover the target genes of more membrane-proximal signaling events? One study of growth factor receptor-mediated signal transduction demonstrated that a largely overlapping repertoire of immediate early genes is modulated by several different growth factor receptors, despite many apparent differences in how the receptors engaged downstream signal transduction proteins[38]. Therefore, it appears that many signaling pathways converge on these immediate early genes, possibly due to their importance in cell cycle progression from G0 to G1. Nonetheless, these growth factor receptors mediate biologically distinct responses, presumably as a result of differential modulation of more delayed response genes. Studies in yeast suggest that individual MAP kinase pathways function as independent signal transduction modules that evoke characteristic changes in target gene expression[39]. Furthermore, perturbations in different components of the same signal transduction pathways in yeast yield similar gene expression changes[40]. Thus, some signal transduction pathways behave in a modular fashion, allowing their contribution to cellular physiology to be separated from other signaling events.

Concluding remarks

The challenge for the future is to generate comprehensive microarray data sets cataloguing the gene expression changes that result from discrete manipulation of individual components of signaling pathways. What is the best way to manipulate signaling pathways to produce interpretable data? Overexpression studies, although easy to perform, can be plagued by secondary effects that result from non-physiological titration of interacting proteins. Thus, when possible, dominant negative, loss-of-function approaches are preferable. Knockout animals can be potentially very useful in probing regulatory pathways.

Acknowledgements
I wish to acknowledge the members of the Staudt laboratory whose unpublished data is summarized in Fig. 2, including Eric Davis, Jena Giltnane, Yunsheng He, Elaine Hurt, Lloyd Lam, Chi Ma, Andreas Rosenwald, Art Shaffer, and Mohan Sathyamoorthy, as well as our collaborators Michael Kuehl, Lief Bersagel, Wyndham Wilson, Nick Chiorazzi, Larry Wahl, Stefan Feske, Anjana Rao, Panu Kovanen, Warren Leonard, Bill Telford, Gerald Marti and Thomas Waldmann. I thank Lance Miller and Ed Liu for their efforts in coordinating the NCI Array facility in which we produce the lymphochip microarrays, and John Powell for designing and maintaining the microarray database at the NCI. I also thank Ash Alizadeh, Mike Eisen, Pat Brown, David Botstein, Research Genetics (especially Troy Moore and Jim Hudson) and the Cancer Genome Anatomy Project (especially Bob Strausberg and Rick Klausner) for help in creating and using the lymphochip microarray.

One important caveat, however, is that if a germ-line mutation leads to altered differentiation of cells, it might be difficult to make a meaningful comparison of gene expression between comparable cell types in wild-type and mutant animals. The use of conditional alleles will be a powerful tool in this regard, both in vivo and in vitro, because they will allow gene expression changes to be studied in a temporal fashion following pathway manipulation. Changes in cellular physiology are often characterized by cascades of gene expression changes caused by serial activation of transcription factors (Fig. 4b). Indeed, if manipulation of a signaling pathway induces cell-cycle arrest, a host of gene expression changes will secondarily occur[39]. To pinpoint the primary gene expression changes downstream of a regulatory protein, conditional alleles can be created by fusing a regulatory protein to the estrogen-receptor ligand-binding domain[25]. Estrogens activate such fusion proteins by inducing a conformational change in the ligand-binding domain and thus the effects of these fusion proteins on gene expression can be studied in the presence of protein synthesis inhibitors. Under these conditions, only the primary targets of the regulatory factor will be affected. Finally, small molecules that discretely modulate signaling protein function will be powerful tools in the near future[41]. This 'chemical genetics' approach will permit modulation of signaling pathways in a temporally defined fashion and could allow multiple regulatory proteins to be modulated in parallel. Pharmacological inhibitors of calcineurin and MEK kinase were tested for their effects on gene expression during B-cell activation using microarrays, and each drug yielded a surprisingly discrete gene expression signature[12]. With suitably powerful mathematical models, well-planned genomic-scale gene expression experiments will lead to a comprehensive understanding of gene expression physiology and, ultimately, to the precise pharmacological manipulation of these pathways to the benefit of patients.

References

1 Schena, M. *et al.* (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10614–10619
2 Lockhart, D.J. *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14, 1675–1680
3 Schena, M. *et al.* (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470
4 Staudt, L.M. and Brown, P.O. (2000) Genomic views of the immune system. *Ann. Rev. Immunol.* 18, 829–859
5 Sherlock, G. (2000) Analysis of large-scale gene expression data. *Curr. Opin. Immunol.* 12, 201–205
6 Eisen, M.B. *et al.* (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14863–14868
7 Alizadeh, A.A. *et al.* (2000) Identification of molecularly and clinically distinct types of diffuse large B-cell lymphoma by gene expression profiling. *Nature* 403, 503–511
8 Alizadeh, A. *et al.* (1999) The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. *Cold Spring Harbor Symp Quant Biol.* 64, 71–78
9 DeRisi, J.L. *et al.* (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278, 680–686
10 Ross, D.T. *et al.* (2000) Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* 24, 227–235
11 Perou, C.M. *et al.* (1999) Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci. U. S. A.* 96, 9212–9217
12 Glynne, R. *et al.* (2000) How self-tolerance and the immunosuppressive drug FK506 prevent B-cell mitogenesis. *Nature* 403, 672–676
13 Teague, T.K. *et al.* (1999) Activation changes the spectrum but not the diversity of genes expressed by T cells. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12691–12696
14 Rogge, L. *et al.* (2000) Transcript imaging of the development of human T helper cells using oligonucleotide arrays. *Nat. Genet.* 25, 96–101
15 Golub, T.R. *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, 531–537
16 Bahler, D.W. and Levy, R. (1992) Clonal evolution of a follicular lymphoma: evidence for antigen selection. *Proc. Natl. Acad. Sci. U. S. A.* 89, 6770–6774
17 Lossos, I.S. *et al.* (2000) Immunoglobulin heavy chain somatic mutation is ongoing in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc. Natl. Acad. Sci.* (in press)
18 Moser, K.L. *et al.* (1998) Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14869–14874
19 Gaffney, P.M. *et al.* (1998) A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14875–14879
20 Zhao, R. et al. (2000) Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. *Genes Dev.* 14, 981–993
21 O'Hagan, R.C. *et al.* (2000) Gene-target recognition among members of the myc superfamily and implications for oncogenesis. *Nat. Genet.* 24, 113–119
22 Coller, H.A. *et al.* (2000) Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proc. Natl. Acad. Sci. U. S. A.* 97, 3260–3265
23 Lee, S.B. *et al.* (1999) The Wilms tumor suppressor WT1 encodes a transcriptional activator of amphiregulin. *Cell* 98, 663–673
24 Khan, J. *et al.* (1999) cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13264–13269
25 Shaffer, A.L. *et al.* (2000) BCL-6 represses genes that function in lymphocyte differentiation, inflammation and cell cycle control. *Immunity* 13, 199–212
26 Staudt, L.M. *et al.* (1999) Regulation of lymphocyte cell fate decisions and lymphomagenesis by BCL-6. *Int. Rev. Immunol.* 18, 381–403
27 Dalla-Favera, R. *et al.* (1999) Molecular pathogenesis of B cell malignancy: the role of BCL-6. *Curr. Top. Microbiol. Immunol.* 246, 257–263
28 Dent, A.L. *et al.* (1997) Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science* 276, 589–592
29 Ye, B.H. *et al.* (1997) The BCL-6 proto-oncogene controls germinal-centre formation and Th2- type inflammation. *Nat. Genet.* 16, 161–170
30 Fukuda, T. *et al.* (1997) Disruption of the Bcl6 gene results in an impaired germinal center formation. *J. Exp. Med.* 186, 439–448
31 Allman, D. *et al.* (1996) BCL-6 expression during B-cell activation. *Blood* 87, 5257–5268
32 Cattoretti, G. *et al.* (1995) Bcl-6 protein is expressed in germinal-center B cells. *Blood* 86, 45–53
33 Onizuka, T. *et al.* (1995) BCL-6 gene product, a 92- to 98-kD nuclear phosphoprotein, is highly expressed in germinal center B cells and their neoplastic counterparts. *Blood* 86, 28–37
34 Lin, Y. *et al.* (1997) Repression of c-myc transcription by Blimp-1, an inducer of terminal B cell differentiation. *Science* 276, 596–599
35 Turner, C.A., Jr. *et al.* (1994) Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell* 77, 297–306
36 Soro, P.G. *et al.* (1999) Differential involvement of the transcription factor Blimp-1 in T cell-independent and -dependent B cell differentiation to plasma cells. *J. Immunol.* 163, 611–617
37 MacLennan, I.C. (1994) Germinal centers. *Annu. Rev. Immunol.* 12, 117–139
38 Fambrough, D. *et al.* (1999) Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell* 97, 727–741
39 Roberts, C.J. *et al.* (2000) Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science* 287, 873–880
40 Hughes, T.R. *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell* 102, 109–126
41 Schreiber, S.L. (2000) Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287, 1964–1969